



УДК 551.515

© 2026 г. Г.В. Гренкин<sup>1</sup>, канд. физ.-мат. наук,  
С.Ю. Глебова<sup>2</sup>, канд. геогр. наук

<sup>1</sup>(Владивостокский государственный университет)

<sup>2</sup>(Тихоокеанский филиал ВНИРО (ТИНРО), Владивосток)

## НЕЧЕТКИЙ МЕТОД ПРОГНОЗИРОВАНИЯ ИЗМЕНЧИВОСТИ КЛИМАТИЧЕСКИХ ПОКАЗАТЕЛЕЙ

Статья посвящена проблеме прогнозирования индексов циркуляции, которые характеризуют ветровой перенос над дальневосточными морями. Поставлена задача предсказать увеличение либо уменьшение временного ряда, если известно значение того же индекса в предыдущий год вместе со значением дополнительного ряда, который выражает причину изменчивости. Предложен нечеткий метод, основанный на оптимальной категоризации входных признаков. Проведено сравнение данного подхода с базовыми методами: линейной регрессией и решающим деревом.

**Ключевые слова:** характеристики атмосферной циркуляции, линейная авторегрессия, решающее дерево, нечетко-множественный подход.

DOI: 10.22250/18142400\_2026\_87\_1\_14

### Введение

Для оценки предполагаемых изменений атмосферных процессов над Дальневосточными морями используются различные показатели. Среди них можно назвать индексы циркуляции А.Л. Каца, показывающие интенсивность зонального и меридионального ветрового переноса над любым выделенным районом, а также состояние центров действия атмосферы. В настоящей работе поставлена задача определить взаимосвязь между характером развития Сибирского антициклона и Алеутской депрессии (местоположение, интенсивность) зимой текущего года с индексами атмосферной циркуляции над Японским морем в зимние месяцы последующего года.

Заблаговременная информация об особенностях ветрового режима над бассейном может представлять интерес для представителей разных областей при составлении путинных прогнозов промысловых объектов. Океанологи используют такие данные для прогноза характера течений, распределения водных масс, сроков начала и окончания развития льда и т.д., и это позволяет ихтиологам давать рекомендации рыбакам о характере предстоящей путины отдельных видов рыб.

Распространенными методами прогнозирования временных рядов являются модели авторегрессии со скользящими средними [1], Фурье-анализ [2]. Эти подходы характерны тем, что не обеспечивают приемлемую точность на длинных временных рядах и требуют пересчета параметров модели при дополнении рядов новыми наблюдениями. В связи с этим существует необходимость в простых и устойчивых инструментах для прогноза динамики небольших временных рядов.

В [3] описан пример применения нечетко-множественного подхода для анализа временных рядов. Чтобы спрогнозировать динамику ряда, в качестве выходного признака принят категориальный признак, обозначающий нечеткий класс приращения значения временного ряда по сравнению с прошлым.

В настоящей работе предпринята попытка описать изменчивость климатических показателей с помощью аналогичной категоризации выходного признака. Поставлена задача предсказания увеличения либо уменьшения временного ряда в зависимости от его значения в предыдущий момент времени, а также значения дополнительного ряда, представляющего причину изменчивости.

В качестве базовых моделей используются методы линейной авторегрессии и решающего дерева. На их основе строятся две нечеткие модификации. Один нечеткий метод устанавливает оптимальное соответствие между диапазонами изменения линейной комбинации двух входных признаков и диапазонами изменения выходного признака на основе критерия минимума энтропии. Другой нечеткий метод осуществляет оптимальную категоризацию входных признаков. Первый из методов аналогичен решающим деревьям с разделяющей гиперплоскостью [4]. Второй представляет собой нечеткий аналог оптимального классифицирующего дерева [5].

### **Материалы и методы**

*Исходные данные.* Расчет индексов А.Л. Каца [6] проводился при помощи синоптических карт 10-дневного осреднения для района 30-50° с.ш., 120-150° в.д. (с шагом 10°), включающего в себя Японское море. Для вычисления зонального индекса ( $I_z$ ) подсчитывалось число изобар, ориентированных с запада на восток и с востока на запад (с учетом циклонической или антициклонической циркуляции воздуха) и пересекающих все меридианы внутри и на границах района; меридиональный индекс ( $I_m$ ) – с севера на юг и с юга на север. Положительные значения  $I_z$  и  $I_m$  характеризуют западный и южный перенос, отрицательные значения – восточный и северный соответственно. Так же при помощи декадных синоптических карт определялась локализация Сибирского антициклона и Алеутской депрессии – сезонных центров действия атмосферы (как статистический результат преобладания в районе барической системы одного знака), который является важным погодным и климатообразующим фактором в регионе.

Данные индексы являются количественным критерием направления и интенсивности «локальных» муссонных переносов над любым выделенным

районом и позволяют оценивать характер многолетней изменчивости ветрового переноса для любого отрезка времени (в нашем случае для января-марта) в течение периода с 1974 по 2023 гг. Методика вычисления индексов циркуляции (с учетом направленности изобар) была описана ранее [7].

*Базовые модели.* Поставим задачу прогнозирования знака приращения временного ряда  $x(t)$

$$z(t+1) = \begin{cases} 1, & \text{если } x(t+1) - x(t) > 0, \\ 0, & \text{иначе} \end{cases}$$

в зависимости от значений  $x(t)$  и  $y(t)$ , где  $y(t)$  – дополнительный ряд, выражающий причину изменчивости первого ряда.

В качестве первой базовой модели используем линейную авторегрессию

$$x(t+1) = ax(t) + by(t) + c$$

с дихотомизацией выходного признака. Сначала вычисляются оптимальные коэффициенты  $a$ ,  $b$ ,  $c$ , обеспечивающие минимум среднеквадратичной ошибки прогноза. Затем находится оптимальный порог  $d$ , обеспечивающий наибольшую точность прогноза при дихотомизации: если  $ax(t) + by(t) \leq d$ , то прогнозируется  $z(t+1) = 0$ , иначе  $z(t+1) = 1$ .

Вторая базовая модель – это решающее дерево глубины 2, с помощью которого предсказывается бинарный выходной признак  $z(t+1)$  с помощью двух входных признаков  $x(t)$ ,  $y(t)$ . В каждом узле дерева задается логическое условие, зависящее от одного из признаков, обеспечивающее наилучшее разделение выборки, исходя из критерия минимальности примеси Джини. В результате получена модель, которую можно представить в виде набора решающих правил такого вида: если  $x(t) > x_0$  и  $y(t) > y_0$ , то  $z(t+1) = z_0$ .

*Нечеткая модель с интегральным показателем.* Под интегральным показателем понимается взвешенная сумма входных признаков  $X(t) = ax(t) + by(t)$ . Разобьем множество возможных значений интегрального показателя на три нечетких класса. Также разобьем множество возможных значений выходного признака  $x(t+1) - x(t)$  на три нечетких класса. Нечеткие множества, которым принадлежит  $x(t+1) - x(t)$ , будут обозначать три возможности: ряд  $x(t)$  возрастает, ряд убывает, либо ряд остается почти неизменным. Поставим задачу подобрать параметры модели, которые обеспечат наилучшее соответствие между диапазонами изменения интегрального показателя и диапазонами выходного признака. В качестве критерия оптимальности будет использован критерий минимума энтропии, выражающий меру определенности, с которой можно определить диапазон выходного признака, зная, какому диапазону принадлежит интегральный показатель.

Введем центроиды диапазонов интегрального показателя – числа  $d_1$ ,  $d_2$ ,  $d_3$ , удовлетворяющие условию  $d_1 \leq d_2 \leq d_3$ . Аналогичным образом опреде-

лим центроиды диапазонов выходного признака — числа  $-\alpha, 0, \alpha$ , где число  $\alpha > 0$  — это небольшая величина выходного признака  $x(t+1) - x(t)$ , при котором можно считать, что ряд  $x(t)$  остается почти неизменным.

Определим меру принадлежности  $s$ -му диапазону интегрального показателя в момент времени  $t$  по формуле

$$u_s(t) = \begin{cases} \frac{X(t) - d_{s-1}}{d_s - d_{s-1}}, & \text{если } X(t) \in [d_{s-1}, d_s], \\ \frac{d_{s+1} - X(t)}{d_{s+1} - d_s}, & \text{если } X(t) \in [d_s, d_{s+1}], \\ 0, & \text{иначе.} \end{cases}$$

Таким образом, мера принадлежности линейно зависит от значения  $X(t)$  на промежутках между серединами диапазонов. Мера принадлежности  $s$ -му диапазону принимает значение 1 в центроиде  $s$ -го диапазона  $d_s$  и значение 0 в серединах соседних диапазонов  $d_{s-1}$  и  $d_{s+1}$ . Полагаем  $d_0 = -\infty, d_4 = +\infty$ . Аналогичным образом определяем меру принадлежности  $p$ -му диапазону выходного признака  $x(t+1) - x(t)$  в момент времени  $t$ , которую обозначим через  $v_p(t)$ . С помощью мер принадлежности определим матрицу соответствия между диапазонами интегрального показателя и диапазонами выходного признака:

$$m_{sp} = \frac{\sum_{t=1}^{T-1} u_s(t) v_p(t)}{\sum_{t=1}^{T-1} u_s(t)}.$$

Здесь  $T$  — длина временных рядов  $x(t), y(t)$ .

Введем целевую функцию, зависящую от весовых коэффициентов  $a, b$  и центроидов  $d_1, d_2, d_3$ , выражающую величину энтропии:

$$J(a, b, d_1, d_2, d_3) = -\frac{1}{T-1} \sum_{t=1}^{T-1} \sum_{s=1}^3 \sum_{p=1}^3 u_s(t) m_{sp} \ln m_{sp}.$$

Поставим задачу оптимизации: требуется найти параметры  $a, b, d_1, d_2, d_3$ , при которых величина целевой функции  $J$  минимальна. Для решения задачи оптимизации применяем метод Нелдера–Мида.

*Нечеткая модель с категоризацией признаков.* Поставим задачу предсказания диапазона выходного признака  $x(t+1) - x(t)$ , если известно, каким диапазонам принадлежат входные признаки  $x(t), y(t)$ . Разобьем множество возможных значений признака  $x(t)$  на 3 класса с центроидами  $d_1, d_2, d_3$ , а множество возможных значений признака  $y(t)$  — на 2 класса с центроидами

$q_1, q_2$ . Обозначим меры принадлежности  $s$ -му диапазону значений  $x(t)$  через  $u_s(t)$ , меры принадлежности  $p$ -му диапазону значений  $y(t)$  – через  $v_p(t)$ , а меры принадлежности  $r$ -му диапазону значений  $x(t+1) - x(t)$  – через  $w_r(t)$ . Меры принадлежности определяются так же, как в предыдущем разделе.

Определим трехмерную матрицу соответствия между диапазонами признака  $x(t)$ , диапазонами признака  $y(t)$  и диапазонами выходного признака  $x(t+1) - x(t)$  по формуле

$$m_{spr} = \frac{\sum_{t=1}^{T-1} u_s(t)v_p(t)w_r(t)}{\sum_{t=1}^{T-1} u_s(t)v_p(t)}.$$

Введем целевую функцию, зависящую от центроидов  $d_1, d_2, d_3$  и  $q_1, q_2$ , выражающую величину энтропии:

$$J(d_1, d_2, d_3, q_1, q_2) = -\frac{1}{T-1} \sum_{t=1}^{T-1} \sum_{s=1}^3 \sum_{p=1}^2 \sum_{r=1}^3 u_s(t)v_p(t)m_{spr} \ln m_{spr}.$$

Поставим задачу оптимизации: требуется найти параметры  $d_1, d_2, d_3, q_1, q_2$ , при которых величина целевой функции  $J$  минимальна. Для решения задачи оптимизации применяем метод Нелдера–Мида.

Применение нечеткого подхода позволяет оптимизировать матрицу соответствия, пользуясь непрерывной целевой функцией. Результатом оптимизации будет набор параметров, обеспечивающий наилучшее соответствие между диапазонами.

### Результаты моделирования

Все описанные методы прогнозирования могут применяться для любых пар временных рядов  $x(t), y(t)$ . В качестве прогнозируемого ряда  $x(t)$  выбираем один из следующих 6 параметров: зональный индекс Iz для Японского моря в январе, феврале, марте; меридиональный индекс Im для Японского моря в январе, феврале, марте. В качестве дополнительного ряда  $y(t)$  выбирались следующие 18 параметров: широта, долгота и интенсивность Сибирского антициклона и Алеутской депрессии в январе, феврале, марте.

Программная реализация выполнена на языке Python [8]. Процедуры построения линейной регрессии и решающего дерева реализованы в пакете scikit-learn, метод оптимизации реализован в пакете scipy.

Применялась нечеткая модель с интегральным показателем для различных пар  $x(t), y(t)$ . В качестве координаты центроида выходного признака принято значение  $\alpha = 0.25$ . Для каждого ряда  $x(t)$  выбран в качестве дополнительного  $y(t)$  тот ряд, для которого оптимум целевой функции  $J$  как можно меньше. В табл. 1 приводится точность прогнозирования.

Таблица 1

$x(t)$	$y(t)$	Линейная авторегрессия	Решающее дерево	Нечеткая модель с интегральным показателем	Нечеткая модель с категоризацией признаков
Iz янв	Сиб инт янв	81.6%	79.5%	81.6%	79.5%
Iz фев	Ал долг фев	75.5%	77.5%	75.5%	77.5%
Iz мар	Сиб шир янв	81.6%	83.6%	81.6%	77.5%
Im янв	Сиб шир фев	81.6%	83.6%	83.6%	81.6%
Im фев	Ал долг мар	77.5%	71.4%	75.5%	73.4%
Im мар	Ал инт фев	81.6%	79.5%	87.7%	83.6%

Рис. 1 иллюстрирует модель линейной авторегрессии для прогнозирования изменчивости временного ряда Iz янв с помощью дополнительного ряда Сиб инт янв. По горизонтальной оси откладываются значения взвешенной суммы входных признаков с параметрами  $a = -1.17$ ,  $b = -0.02$ . По вертикальной оси откладываются значения выходного признака. Вертикальная прямая на рисунке обозначает пороговое значение взвешенной суммы признаков, выше которого прогнозируется возрастание временного ряда.

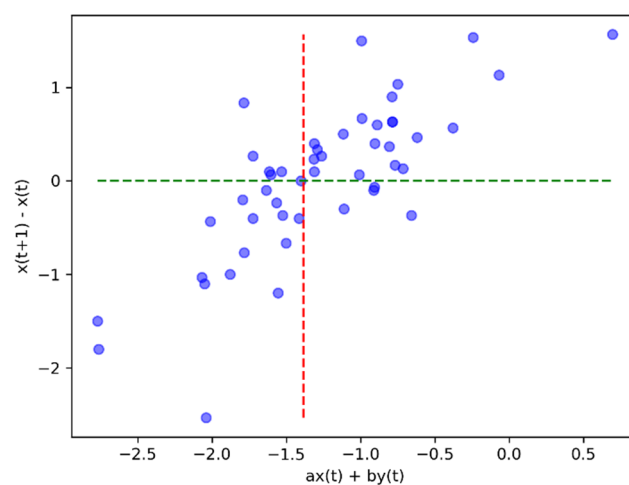


Рис. 1. Пример применения модели линейной авторегрессии.

В табл. 2 представлены правила, к которым сводится модель решающего дерева.

Таблица 2

Условие	Прогноз	Охват выборки	Точность прогноза
$x(t) \leq 0.383$ и $y(t) \leq 49.0$	возрастание	49%	88%
$x(t) \leq 0.383$ и $y(t) > 49.0$	убывание	2%	100%
$0.383 < x(t) \leq 0.817$	убывание	35%	59%
$x(t) > 0.817$	убывание	14%	100%

На рис. 2 представлен результат применения нечеткой модели с интегральным показателем для прогнозирования изменчивости временного ряда

Из янв с помощью дополнительного ряда Сиб инт янв. Каждое наблюдение отмечено знаком «+» либо «-», обозначающим результат прогнозирования. Рядом указан символ «!», если прогноз ошибочен.

Найдены следующие параметры модели:  $a = -1.24$ ,  $b = -0.02$ ,  $d_1 = -1.45$ ,  $d_2 = -1.36$ ,  $d_3 = 0.77$ . Оптимизированная матрица соответствия  $m_{sp}$

имеет вид: 
$$\begin{pmatrix} 0.68 & 0.17 & 0.13 \\ 0.10 & 0.14 & 0.75 \\ 0.08 & 0.09 & 0.82 \end{pmatrix}.$$

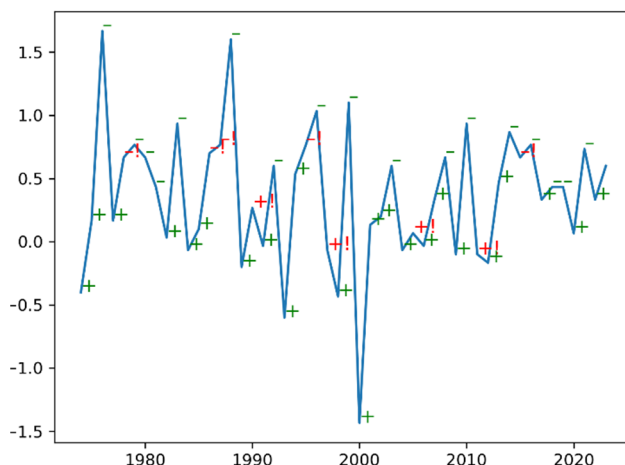


Рис. 2. Пример прогнозирования изменчивости временного ряда.

В табл. 3 представлена оптимизированная матрица соответствия, построенная по нечеткой модели с категоризацией признаков. Строки таблицы соответствуют диапазонам изменения признака  $x(t)$ , столбцы соответствуют диапазонам изменения выходного признака  $x(t+1) - x(t)$ . В ячейках таблицы записаны два значения  $m_{s1r}$ ,  $m_{s2r}$ , соответствующие двум диапазонам изменения признака  $y(t)$ . Найдены значения параметров:  $d_1 = -1.19$ ,  $d_2 = 0.12$ ,  $d_3 = 0.43$ ,  $q_1 = 48$ ,  $q_2 = 49$ .

Таблица 3

	$r = 1$	$r = 2$	$r = 3$
$s = 1$	4%   26%	3%   73%	92%   0%
$s = 2$	10%   26%	12%   73%	77%   0%
$s = 3$	52%   100%	18%   0%	30%   0%

На рассмотренном примере линейная авторегрессия и нечеткая модель с интегральным показателем показали одинаковую точность. Модель решающего дерева сгенерировала тривиальное решающее правило, в котором дополнительный ряд  $y(t)$  практически не задействуется, а возрастание ряда  $x(t)$  предсказывается тогда, когда предшествующее значение  $x(t)$  не превосходит порогового значения. Тем не менее, даже это простое решающее правило обеспечивает высокую точность прогноза.

Наиболее понятной и интерпретируемой моделью является решающее дерево, которое показало в целом меньшую точность, чем модели, использующие линейную комбинацию входных признаков. Результаты демонстри-

руют, что введение в модель дополнительного ряда  $y(t)$  может незначительно повысить точность прогнозирования. Таким образом, оптимальной моделью для прогнозирования изменчивости индексов атмосферной циркуляции остается линейная авторегрессия и ее модификации.

Однако, как видно из табл. 1, в случаях, когда  $y(t)$  представляет собой ряд интенсивности Алеутской депрессии или Сибирского антициклона, лучше работает модель линейной авторегрессии. В случаях же, когда  $y(t)$  представляет собой ряд широты или долготы, лучше справляется модель решающего дерева. Можно предположить, что влияние интенсивности на индексы циркуляции происходит в большей степени линейно, а влияние положения центра действия атмосферы – скачкообразно.

### Заключение

В статье рассмотрена проблема прогнозирования климатических показателей по дополнительному ряду, выражающему причинность изменчивости. Для решения данной задачи предложено два нечетких метода с категоризацией интегрального показателя и категоризацией отдельных признаков. Первый метод позволил незначительно повысить точность прогнозирования по сравнению с базовой моделью линейной авторегрессии.

Дальнейшая работа может быть направлена на достижение компромисса между точностью и простотой модели за счет использования множественной кусочно-линейной регрессии.

### ЛИТЕРАТУРА

1. Айвазян С.А. Прикладная статистика. Основы эконометрики: Учебник для вузов: Т. 2: Основы эконометрики. – М.: ЮНИТИ-ДАНА, 2001.
2. Цицаишвили Г.Ш., Шатилина Т.А., Радченкова Т.В. Применение новых алгоритмов обработки метеорологических наблюдений: монография. – Казань: Бук, 2022.
3. Ярушкина Н.Г., Афанасьева Т.В., Перфильева И.Г. Интеллектуальный анализ временных рядов: учебное пособие. – Москва: ИД «ФОРУМ»: ИНФРА-М, 2022.
4. Luna-Benoso B. et al. A new classification model using a decision tree generated from hyperplanes in dimensional space // Applied Artificial Intelligence. – 2024. – Vol. 38. – No. 1. – p. 2426377.
5. Bertsimas D., Dunn J. Optimal classification trees // Machine Learning. – 2017. – Vol. 106. – pp. 1039–1082.
6. Кац А.Л. Сезонные изменения общей циркуляции атмосферы и долгосрочные прогнозы. – Ленинград: Гидрометеиздат, 1960.
7. Глебова С.Ю. Особенности развития атмосферных процессов над Охотским морем в 2000-2006 гг. // Известия ТИНРО. – 2007. – Т. 150. – С. 200–216.
8. <https://github.com/lapkin25/dynamic-regression> (дата обращения: 16.06.2025).

*Статья представлена к публикации членом редколлегии Л.В. Никифоровой.*

*E-mail:*

*Гренкин Глеб Владимирович – [Gleb.Grenkin@vvsu.ru](mailto:Gleb.Grenkin@vvsu.ru);*

*Глебова Светлана Юрьевна – [svetlana.glebova@tinro.vniro.ru](mailto:svetlana.glebova@tinro.vniro.ru).*